ELSEVIER

Bull. Sci. math. 135 (2011) 775–794

BULLETIN DES SCIENCES MATHÉMATIQUES

www.elsevier.com/locate/bulsci

Fokker–Planck equation with respect to heat measures on loop groups

Shizan Fang [a,1], Jinghai Shao [b,*,2]

[a] IMB, BP 47870, Université de Bourgogne, Dijon, France
[b] School of Mathematical Sciences, Beijing Normal University, Beijing, China

Available online 6 July 2011

In Memory of Paul Malliavin

Abstract

The Dirichlet form on the loop group $\mathcal{L}_e(G)$ with respect to the heat measure defines a *Laplacian* Δ^{DM} on $\mathcal{L}_e(G)$. In this note, we will use Wasserstein distance variational method to solve the associated heat equation for a given data of finite entropy.

1. Introduction

Malliavin calculus gives a profound link between Hörmander's bracket phenomenon and the smoothness of the law of the associated diffusion [12]. At the beginning of 1990, Malliavin was interested in quasi-invariance of measures on infinite-dimensional topological non-commutative groups and wrote a series of papers on loop groups [13–16]. In [16], he introduced heat measures on loop groups; these measures have good behaviors such as quasi-invariance [3], functional inequalities [4,5,22] and Monge–Kantorovich optimal transportation [6].

Now let us explain more precisely the content of this paper. Let G be a connected compact Lie group and $\mathcal{G}$ its Lie algebra equipped with Ad_G-invariant inner product $\langle,\rangle$. Consider the based loop groups $\mathcal{L}_e(G)$:

$$\mathcal{L}_e(G) := \big\{\ell : [0,1] \to G \text{ continuous; } \ell(0) = \ell(1) = e\big\},$$

* Corresponding author.
E-mail address: jinghai.shao@yahoo.com.cn (J. Shao).

1 This work has been benefited from financial support: ANR-09-BLAN-0364-01.
2 Benefited from financial support: FANEDD (200917).

0007-4497/$ – see front matter
doi:10.1016/j.bulsci.2011.07.004

where e is the unit element of G. Let $H_0(\mathcal{G})$ be the Cameron–Martin space of absolutely continuous curves h from $[0, 1]$ to $\mathcal{G}$ such that $h(0) = h(1) = 0$ and $|h|_{H_0}^2 = \int_0^1 |\dot{h}(\theta)|_{\mathcal{G}}^2 \, d\theta < +\infty$. The space $H_0(\mathcal{G})$, equipped with the Lie bracket: $[h_1, h_2](\theta) := [h_1(\theta), h_2(\theta)]$, plays the role of the Lie algebra of $\mathcal{L}_e(G)$.

Let

$$W_0(\mathcal{G}) = \bigl\{w : [0, 1] \to \mathcal{G} \text{ continuous; } w(0) = w(1) = 0\bigr\}.$$

Then $(W_0(\mathcal{G}), H_0(\mathcal{G}))$ together with the Brownian bridge measure μ_0 on $W_0(\mathcal{G})$ is an abstract Wiener space. Let $x(t, \cdot)$ be a Brownian motion taking values on $W_0(\mathcal{G})$, with the covariance operator $\langle , \rangle_{H_0}$. For each $\theta \in [0, 1]$, we consider the s.d.e. on G

$$d_t g_x(t, \theta) = g_x(t, \theta) \circ d_t x(t, \theta), \qquad g_x(0, \theta) = e, \tag{1.1}$$

where d_t denotes the Stratonovich stochastic differential relative to the time t. It was proved in [16,3] that $(t, \theta) \mapsto g_x(t, \theta)$ admits a continuous version, that we denote by the same notation. Then we get a continuous stochastic process $t \mapsto g_x(t, \cdot)$ on $\mathcal{L}_e(G)$. Let ν denote the law of $x \mapsto g_x(1, \cdot)$ on $\mathcal{L}_e(G)$, which is called the *heat measure* on $\mathcal{L}_e(G)$.

A function $F : \mathcal{L}_e(G) \to \mathbb{R}$ is said to be cylindrical if there exists a finite partition $\mathcal{P} = \{0 < \theta_1 < \cdots < \theta_n < 1\}$ and a function $f \in C^\infty(G^n)$ such that

$$F(\ell) = f\bigl(\ell(\theta_1), \ldots, \ell(\theta_n)\bigr), \quad \forall \ell \in \mathcal{L}_e(G).$$

The totality of cylindrical functions on $\mathcal{L}_e(G)$ is denoted by $\mathbf{Cylin}(\mathcal{L}_e(G))$. For a cylindrical function F in above form and $h \in H_0(\mathcal{G})$, we define

$$D_h F(\ell) = \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} F\bigl(\ell e^{\varepsilon h}\bigr) = \sum_{i=1}^{n} \bigl\langle \ell^{-1}(\theta_i)\partial_i f, h(\theta_i)\bigr\rangle_{\mathcal{G}},$$

where $\partial_i f$ is the ith partial derivative of f. We say that the gradient of F exists if there exists a $H_0(\mathcal{G})$-valued Z such that for each $h \in H_0(\mathcal{G})$: $D_h F(\ell) = \langle Z(\ell), h\rangle_{H_0}$. Denote the gradient of F by $\nabla^L F$. For cylindrical function F,

$$\bigl(\nabla^L F\bigr)(\ell) = \sum_{i=1}^{n} \ell^{-1}(\theta_i)(\partial_i f) G(\theta_i, \cdot), \tag{1.2}$$

where $G(\theta_i, \theta)$ is the Green function on the circle:

$$G(\theta_i, \theta) = \theta_i \wedge \theta - \theta_i \theta. \tag{1.3}$$

The following result of quasi-invariance was arisen in [16] and proved in [3].

Theorem 1.1. *Let $h \in H_0(\mathcal{G})$. Then there exists K_h in all $L^p(\mathcal{L}_e(G), \nu)$ such that*

$$\int_{\mathcal{L}_e(G)} D_h F(\ell) \, d\nu(\ell) = \int_{\mathcal{G}} F(\ell) K_h(\ell) \, d\nu(\ell). \tag{1.4}$$

By (1.2) and (1.4), we see that the quadratic form $\int_{\mathcal{L}_e(G)} |\nabla^L F|_{H_0}^2 \, d\nu$ is closable on $\mathbf{Cylin}(\mathcal{L}_e(G))$ and

$$\int_{\mathcal{L}_e(G)} \bigl|\nabla^L F\bigr|_{H_0}^2 \, d\nu = \int_{\mathcal{L}_e(G)} \Delta^{DM} F \cdot F \, d\nu, \tag{1.5}$$

where Δ^{DM} is the associated infinitesimal generator. By theory of Dirichlet form [17], the semigroup $e^{-t\Delta^{DM}} F$ is well defined for $F \in L^p(\mathcal{L}_e(G), \nu)$ with $p > 1$.

Before stating the main result of this note, let's introduce some notation. We will denote by $\mathcal{P}(\mathcal{L}_e(G))$ the space of probability measures on $\mathcal{L}_e(G)$. A continuous curve $\gamma : [0,1] \to \mathcal{L}_e(G)$ is said to be *admissible* if there exists $Z_t = \int_0^t Z'_s \, ds \in H_0(\mathcal{G})$ with $\int_0^1 |Z'_s|^2_{H_0} \, ds < +\infty$ such that for $\theta \in [0,1]$,

$$d_t \gamma(t,\theta) = \gamma(t,\theta) Z'_t(\theta) \, dt, \qquad \gamma(0,\theta) = e, \tag{1.6}$$

where d_t denotes the derivative w.r.t. t. For an admissible curve γ, we define its length $L(\gamma)$ by

$$L(\gamma) = \left(\int_0^1 \left| Z'_t \right|^2_{H_0} dt \right)^{1/2}.$$

For a non-admissible curve, its length is defined to be $+\infty$. The *Riemannian distance* d_L on $\mathcal{L}_e(G)$ is defined as, for any two points $\ell_1, \ell_2 \in \mathcal{L}_e(G)$,

$$d_L(\ell_1, \ell_2) = \inf\left\{ L(\gamma);\ \gamma \text{ continuous curve connecting } \mathbf{e} \text{ and } \ell_1^{-1}\ell_2 \right\}, \tag{1.7}$$

where $\mathbf{e}$ denotes the identity loop. It is clear that d_L is left invariant: $d_L(\ell_1, \ell_2) = d_L(\mathbf{e}, \ell_1^{-1}\ell_2)$ for any $\ell_1,\ \ell_2 \in \mathcal{L}_e(G)$. It was proved in [6] that $(\ell_1, \ell_2) \mapsto d_L(\ell_1, \ell_2)$ is lower semi-continuous from $\mathcal{L}_e(G) \times \mathcal{L}_e(G)$ to $[0, +\infty]$.

For two probability measures μ_1 and μ_2 on $\mathcal{L}_e(G)$, we define the Wasserstein distance between them by

$$W_2(\mu_1, \mu_2) = \inf_{\pi \in \mathcal{C}(\mu_1, \mu_2)} \left(\int_{\mathcal{L}_e(G) \times \mathcal{L}_e(G)} d_L(\ell_1, \ell_2)^2 \, d\pi(\ell_1, \ell_2) \right)^{1/2}, \tag{1.8}$$

where $\mathcal{C}(\mu_1, \mu_2)$ denotes the collection of all probability measures on $\mathcal{L}_e(G) \times \mathcal{L}_e(G)$ with marginals μ_1 and μ_2 respectively.

For $\mu \in \mathcal{P}(\mathcal{L}_e(G))$ which admits a density ρ with respect to ν, we denote by $\mathbf{Ent}_\nu(\mu)$ the entropy of μ with respect to ν:

$$\mathbf{Ent}_\nu(\mu) = \int_{\mathcal{L}_e(G)} \rho \log \rho \, d\nu.$$

Theorem 1.2. *Let $\mu_0 \in \mathcal{P}(\mathcal{L}_e(G))$ of finite entropy $\mathbf{Ent}_\nu(\mu_0) < +\infty$; then there exists a unique continuous curve $\mu : [0,1] \to \mathcal{P}(\mathcal{L}_e(G))$, which solves*

$$\int_{[0,1] \times \mathcal{L}_e(G)} \left[\alpha'(t) F(\ell) - \alpha(t) \Delta^{DM} F(\ell) \right] d\mu_t(\ell) \, dt = \alpha(0) \int_{\mathcal{L}_e(G)} F \, d\mu_0, \tag{1.9}$$

for all $\alpha \in C_c^\infty([0,1[)$, and $F \in \mathbf{Cylin}(\mathcal{L}_e(G))$. Moreover

$$W_2(\mu_t, \mu_s) \leqslant \sqrt{6\, \mathbf{Ent}_\nu(\mu_0)} \sqrt{|t-s|}.$$

In other words, μ_t solves the heat equation in weak sense:

$$\frac{d\mu_t}{dt} = -\Delta^{DM} \mu_t. \tag{1.10}$$

2. Finite-dimensional approximation (see [4])

The Wasserstein distance variational formulation for Fokker–Planck equations has been introduced in [10] (see also [20]). Gradient flow on some infinite-dimensional situations has been discussed in [8,2,7]. Following [21,11,18,23,24], the convexity of the functional $\mathbf{Ent}_\nu$ is related to lower boundedness of the Ricci curvature associated to Dirichlet form (1.5), that is not available. It seems that the Fokker–Planck equation is a weaker version of gradient flows.

Let $\mathcal{P} = \{0 < \theta_1 < \cdots < \theta_n < 1\}$ be a finite partition of [0, 1]. For any $h \in H_0(\mathcal{G})$ we define

$$\Pi_{\mathcal{P}} h = \sum_{ij} G(\theta_i, \cdot) Q_{ij}^{\mathcal{P}} h(\theta_j), \tag{2.1}$$

where $(Q_{ij}^{\mathcal{P}})$ is the inverse matrix of $(G(\theta_i, \theta_j))_{1 \leqslant i,j \leqslant n}$, and $G(\theta_i, \theta_j)$ was defined in (1.3). Note that $(\Pi_{\mathcal{P}} h)(\theta_i) = h(\theta_i)$ for $1 \leqslant i \leqslant n$. Set

$$H_{\mathcal{P}}(\mathcal{G}) = \big\{\Pi_{\mathcal{P}} h;\ h \in H_0(\mathcal{G})\big\}. \tag{2.2}$$

It's easy to verify that $\langle \Pi_{\mathcal{P}} h, k\rangle_{H_0} = \langle \Pi_{\mathcal{P}} h, \Pi_{\mathcal{P}} k\rangle_{H_0}$ for any $h, k \in H_0(\mathcal{G})$, so $\Pi_{\mathcal{P}}$ is an orthogonal projection from $H_0(\mathcal{G})$ onto $H_{\mathcal{P}}(\mathcal{G})$. Define $\Lambda_{\mathcal{P}} : H_0(\mathcal{G}) \to \mathcal{G}^{\mathcal{P}}$ by

$$\Lambda_{\mathcal{P}}(h) = \big(h(\theta_1), \ldots, h(\theta_n)\big). \tag{2.3}$$

Then $\Lambda_{\mathcal{P}}$ is an isometric isomorphism from $H_{\mathcal{P}}(\mathcal{G})$ onto $\mathcal{G}^{\mathcal{P}}$ if $\mathcal{G}^{\mathcal{P}}$ is equipped with the inner product defined by

$$\langle a, b\rangle_{\mathcal{P}} := \sum_{i,j=1}^{n} Q_{ij}^{\mathcal{P}} \langle a_i, b_j\rangle_{\mathcal{G}}, \quad a = (a_1, \ldots, a_n),\ b = (b_1, \ldots, b_n) \in \mathcal{G}^n. \tag{2.4}$$

We denote its inverse by $\Lambda_{\mathcal{P}}^{-1} : \mathcal{G}^{\mathcal{P}} \to H_{\mathcal{P}}(\mathcal{G})$. For a cylindrical function F in the form $F = f \circ \Lambda_{\mathcal{P}}$, it is easy to check that for any $h \in H_0(\mathcal{G})$,

$$\big\langle \nabla^L F, h\big\rangle_{H_0} = \big\langle \nabla^L F, \Pi_{\mathcal{P}} h\big\rangle_{H_0}, \tag{2.5}$$

which implies that $\nabla^L F \in H_{\mathcal{P}}(\mathcal{G})$. For $a = (a_1, \ldots, a_n) \in \mathcal{G}^n$, set

$$\tilde{a} = \sum_{ij} G(\theta_i, \cdot) Q_{ij}^{\mathcal{P}} a_j. \tag{2.6}$$

It has been calculated in [3] that

$$\big\langle \nabla^L F(\ell), \tilde{a}\big\rangle_{H_0} = \big\langle \nabla f\big(\Lambda_{\mathcal{P}}(\ell)\big), a\big\rangle_{\mathcal{P}} = \sum_i \big\langle \ell(\theta_i)^{-1} (\partial_i f), a_i\big\rangle_{\mathcal{G}}. \tag{2.7}$$

It has been shown in [3] that the probability measure $\nu_{\mathcal{P}} := (\Lambda_{\mathcal{P}})_* \nu$ on $G^{\mathcal{P}}$ is the heat kernel measure when its Lie algebra $\mathcal{G}^{\mathcal{P}}$ is equipped with the inner product $\langle\,,\rangle_{\mathcal{P}}$ defined in (2.4). As a product space, $(G^{\mathcal{P}}, \langle\,,\rangle_{\mathcal{P}})$ is also a connected compact Lie group. We will denote by $d_{\mathcal{P}}(\cdot,\cdot)$ the associated Riemannian distance on $G^{\mathcal{P}}$. The following result is taken from [6]

Proposition 2.1. *Let $(\mathcal{P}_n)_{n \geqslant 1}$ be a sequence of partition of* [0, 1] *such that $\mathcal{P}_n \subset \mathcal{P}_{n+1}$ and $\bigcup_{n \geqslant 1} \mathcal{P}_n$ is dense in* [0, 1]. *Then for any $\ell_1, \ell_2 \in \mathcal{L}_e(G)$,*

$$d_{\mathcal{P}_n}(\Lambda_{\mathcal{P}_n} \ell_1, \Lambda_{\mathcal{P}_n} \ell_2) \uparrow d_L(\ell_1, \ell_2). \tag{2.8}$$

Given a finite partition $\mathcal{P}$ of [0, 1], the projection of ν to $G^{\mathcal{P}}$ by $\Lambda_{\mathcal{P}}$ is the heat kernel measure $\nu_{\mathcal{P}}$ on $G^{\mathcal{P}}$, whose metric is induced by the inner product $\langle,\rangle_{\mathcal{P}}$ (see (2.4)) on $\mathcal{G}^{\mathcal{P}}$. Let div_ν (resp. $\mathrm{div}_{\nu_{\mathcal{P}}}$) denote the divergence operator on $\mathcal{L}_e(G)$ with respect to ν (resp. $\mathrm{div}_{\nu_{\mathcal{P}}}$ on $G^{\mathcal{P}}$), that is, for $Z : \mathcal{L}_e(G) \to H_0(\mathcal{G})$ a vector field,

$$\int_{\mathcal{L}_e(G)} \big\langle \nabla F(\ell), Z(\ell)\big\rangle_{H_0}\, d\nu(\ell) = \int_{\mathcal{L}_e(G)} F(\ell)\,\mathrm{div}_\nu(Z)(\ell)\, d\nu(\ell). \tag{2.9}$$

Set $\mathcal{B} = \mathcal{B}(G^{\mathcal{P}})$ the Borel σ-field of $G^{\mathcal{P}}$, and $\mathbb{E}_\nu^{\mathcal{B}}(Z)$ the conditional expectation of Z with respect to $\mathcal{B}$, that is, $x \mapsto (E_\nu^{\mathcal{B}} Z)(x)$ is $\mathcal{B}$ measurable, and $\forall f \in C^\infty(G^{\mathcal{P}})$, $\forall h \in H_0(\mathcal{G})$

$$\int_{G^{\mathcal{P}}} f(x)\big\langle h, \big(\mathbb{E}_\nu^{\mathcal{B}} Z\big)(x)\big\rangle_{H_0}\, d\nu_{\mathcal{P}}(x) = \int_{\mathcal{L}_e(G)} f \circ \Lambda_{\mathcal{P}}(\ell)\big\langle h, Z(\ell)\big\rangle_{H_0}\, d\nu(\ell). \tag{2.10}$$

Proposition 2.2. *Assume that* $\mathrm{div}_\nu(Z) \in L^2(\nu)$ *exists, then*

$$\mathrm{div}_{\mathcal{P}}(Z_{\mathcal{P}}) = \mathbb{E}_\nu^{\mathcal{B}}\big(\mathrm{div}_\nu(\widetilde{\Lambda_{\mathcal{P}} Z})\big), \tag{2.11}$$

where $Z_{\mathcal{P}} = \mathbb{E}_\nu^{\mathcal{B}}(\Lambda_{\mathcal{P}} Z)$ *and the meaning of* $\sim$ *in* (2.11) *is given in* (2.6).

Proof. For any $f \in C^\infty(G^{\mathcal{P}})$, we have

$$\begin{aligned}\int_{G^{\mathcal{P}}} f\,\mathrm{div}_{\nu_{\mathcal{P}}}(Z_{\mathcal{P}})\, d\nu_{\mathcal{P}} &= \int_{G^{\mathcal{P}}} \langle \nabla f, Z_{\mathcal{P}}\rangle_{\mathcal{P}}\, d\nu_{\mathcal{P}} \\ &= \int_{G^{\mathcal{P}}} \big\langle \nabla f, \mathbb{E}_\nu^{\mathcal{B}}(\Lambda_{\mathcal{P}} Z)\big\rangle_{\mathcal{P}}\, d\nu_{\mathcal{P}} = \int_{\mathcal{L}_e(G)} \langle \nabla f \circ \Lambda_{\mathcal{P}}, \Lambda_{\mathcal{P}} Z\rangle_{\mathcal{P}}\, d\nu.\end{aligned}$$

By (2.7), the above quantity yields

$$\begin{aligned}\int_{\mathcal{L}_e(G)} \big\langle \nabla^{\mathcal{L}}(f \circ \Lambda_{\mathcal{P}}), \widetilde{\Lambda_{\mathcal{P}} Z}\big\rangle_{H_0}\, d\nu &= \int_{\mathcal{L}_e(G)} f \circ \Lambda_{\mathcal{P}}\,\mathrm{div}_\nu(\widetilde{\Lambda_{\mathcal{P}} Z})\, d\nu \\ &= \int_{G^{\mathcal{P}}} f\,\mathbb{E}_\nu^{\mathcal{B}}\big(\mathrm{div}_\nu(\widetilde{\Lambda_{\mathcal{P}} Z})\big)\, d\nu_{\mathcal{P}}.\end{aligned}$$

The result (2.11) follows. □

In particular, applying (2.11) to $Z = \nabla^{\mathcal{L}} F$, for $F \in \mathbf{Cylin}(\mathcal{L}_e(G))$, we get

$$\mathrm{div}_{\nu_{\mathcal{P}}}(Z_{\mathcal{P}}) = \mathbb{E}_\nu^{\mathcal{B}}\big(\Delta^{DM} F\big). \tag{2.12}$$

Indeed, we have by (1.2),

$$Z(\ell) = \nabla^{\mathcal{L}} F(\ell) = \sum_{i=1}^n \ell(\theta_i)^{-1}\partial_i f\big(\Lambda_{\mathcal{P}}(\ell)\big) G(\theta_i, \cdot);$$

therefore the kth component of $\Lambda_{\mathcal{P}} Z \in \mathcal{G}^{\mathcal{P}}$ has the expression

$$\sum_{i=1}^n \ell(\theta_i)^{-1}\partial_i f \circ \Lambda_{\mathcal{P}} G(\theta_i, \theta_k).$$

Hence, for $x = (x_1, \dots, x_n) \in G^{\mathcal{P}}$,

$$Z_{\mathcal{P}}(x) = \left(\sum_{i=1}^n x_i^{-1} \partial_i f(x) G(\theta_i, \theta_1), \dots, \sum_{i=1}^n x_i^{-1} \partial_i f(x) G(\theta_i, \theta_n) \right).$$

According to the definition (2.6),

$$\begin{aligned} \widetilde{\Lambda_{\mathcal{P}} Z}(\ell) &= \sum_{j,k=1}^n G(\theta_j, \cdot) Q_{jk}^{\mathcal{P}} \sum_{i=1}^n \ell(\theta_i)^{-1} \partial_i f\big(\Lambda_{\mathcal{P}}(\ell)\big) G(\theta_i, \theta_k) \\ &= \sum_{i=1}^n G(\theta_i, \cdot) \ell(\theta_i)^{-1} \partial_i f\big(\Lambda_{\mathcal{P}}(\ell)\big) = \nabla^L F(\ell). \end{aligned}$$

Now using (2.11), we get (2.12). Furthermore, we have

$$\big\langle Z_{\mathcal{P}}(x), a \big\rangle_{\mathcal{P}} = \sum_{i=1}^n \big\langle x_i^{-1} \partial_i f, a_i \big\rangle_{\mathcal{G}} = D_a f(x).$$

It follows that $Z_{\mathcal{P}} = \nabla f$.

3. Proof of Theorem 1.2

The direct approach via Wasserstein distance variational method will meet two difficulties: 1. The formula of the change of variables: the quasi-invariance for the heat measure ν on $\mathcal{L}_e(G)$ has been established, only for left invariant vector fields [3]; 2. The existence of Monge optimal transport maps has been established only when the initial measure is ν [6]. We will use finite-dimensional approximations. But first of all, we will prepare some results of compactness.

Proposition 3.1. *Let K be a compact set in $\mathcal{L}_e(G)$ with respect to the uniform topology, then for each $R > 0$ the set $K_R := \{\ell \in \mathcal{L}_e(G);\ d_L(\ell, K) \leqslant R\}$ is also compact, where*

$$d_L(\ell, K) = \inf\big\{ d_L\big(\ell, \ell'\big);\ \ell' \in K \big\}.$$

Proof. Let's first show that the closed ball $B_R := \{\ell \in \mathcal{L}_e(G);\ d_L(\mathbf{e}, \ell) \leqslant R\}$ is compact. Let $\ell_n \in B_R$, $n \in \mathbb{N}$. Then by definition of d_L, there exists an admissible curve $\gamma_n : [0, 1] \to \mathcal{L}_e(G)$ and $Z_n \in H(H_0(\mathcal{G}))$ such that

$$d_t \gamma_n(t, \theta) = \gamma_n(t, \theta) Z_n'(t, \theta)\, dt, \qquad \gamma_n(0, \theta) = e, \qquad \gamma_n(1, \theta) = \ell_n(\theta),$$

and

$$d_L(\mathbf{e}, \ell_n) \geqslant L(\gamma_n) - \frac{1}{n}.$$

We have, therefore,

$$\int_0^1 \big| Z_n'(t) \big|_{H_0}^2\, dt \leqslant R^2 + \frac{1}{n} \leqslant R^2 + 1.$$

So, up to a subsequence, Z_n converges weakly to some $Z \in H(H_0(\mathcal{G}))$ and

$$\left(\int_0^1 \left|Z'(t)\right|_{H_0}^2 dt\right)^{1/2} \leqslant \liminf_{n\to\infty}\left(\int_0^1 \left|Z_n'(t)\right|_{H_0}^2 dt\right)^{1/2} \leqslant R. \tag{3.1}$$

Define an admissible curve $\gamma : [0,1] \to \mathcal{L}_e(G)$ by

$$d_t\gamma(t,\theta) = \gamma(t,\theta)Z'(t,\theta)\,dt, \qquad \gamma(0,\theta) = e, \quad \theta \in [0,1].$$

Then γ_n converges pointwisely to γ. Moreover, it is easy to see $\{\gamma_n;\ n \geqslant 1\}$ is an equi-continuous family on $[0,1]\times[0,1]$, so γ_n converges uniformly to γ. In particular, $\ell_n = \gamma_n(1)$ converges uniformly to $\ell := \gamma(1) \in \mathcal{L}_e(G)$. Since $d_L(\mathbf{e}, \ell) \leqslant L(\gamma) = (\int_0^1 |Z'(t)|_{H_0}^2\,dt)^{1/2} \leqslant R$, we get $\ell \in B_R$ and this shows that B_R is compact. Now we consider the set K_R. Let $\tilde{\ell}_n \in K_R$. There exists $k_n \in K$ such that $d_L(\tilde{\ell}_n, k_n) \leqslant R + \frac{1}{n}$. Up to a subsequence, k_n converges to $k \in K$. By left invariance of d_L, $d_L(\mathbf{e}, \tilde{\ell}_n^{-1}k_n) = d_L(\tilde{\ell}_n, k_n) \leqslant R + \frac{1}{n} \leqslant R + 1$. Then there exists a subsequence n_m such that $\tilde{\ell}_{n_m}^{-1}k_{n_m}$ converges to $\bar{\ell}$ in $\mathcal{L}_e(G)$. Then $\tilde{\ell}_{n_m} = k_{n_m}(\tilde{\ell}_{n_m}^{-1}k_{n_m})^{-1}$ converges to some $\ell \in \mathcal{L}_e(G)$. The lower semi-continuity of d_L yields

$$d_L(\ell, k) \leqslant \liminf_{m\to\infty} d_L(\tilde{\ell}_{n_m}, k_{n_m}) \leqslant R.$$

So $\ell \in K_R$, which implies that K_R is compact. $\square$

Proposition 3.2. *For each $R > 0$ and each fixed probability measure σ_0 on $\mathcal{L}_e(G)$, the set*

$$B_R(\sigma_0) := \left\{\mu \in \mathcal{P}\big(\mathcal{L}_e(G)\big);\ W_2(\mu, \sigma_0) \leqslant R\right\}$$

is compact in $\mathcal{P}(\mathcal{L}_e(G))$ with respect to the weak convergence topology.

Proof. Let $\mu_n \in B_R(\sigma_0)$, $n \in \mathbb{N}$. Then for each n there exists $\pi_n \in C(\mu_n, \sigma_0)$ such that

$$W_2(\mu_n, \sigma_0) = \left(\int_{\mathcal{L}_e(G)\times\mathcal{L}_e(G)} d_L(\ell_1, \ell_2)^2\, d\pi_n(\ell_1, \ell_2)\right)^{1/2} \leqslant R.$$

Let $\varepsilon > 0$, there exists a compact set $K \subset \mathcal{L}_e(G)$ such that $\sigma_0(K) \geqslant 1 - \varepsilon$. Let

$$K_r = \left\{\ell \in \mathcal{L}_e(G);\ d_L(\ell, K) \leqslant r\right\}, \quad r > 0,$$

and K_r^c denote its complement in $\mathcal{L}_e(G)$. Proposition 3.1 says that K_r is compact. We have

$$\begin{aligned}
\mu_n\big(K_r^c\big) &= \int_{\mathcal{L}_e(G)\times\mathcal{L}_e(G)} \mathbf{1}_{K_r^c}(\ell_1)\, d\pi_n(\ell_1, \ell_2) \\
&= \int_{\mathcal{L}_e(G)\times K} \mathbf{1}_{K_r^c}(\ell_1)\, d\pi_n(\ell_1, \ell_2) + \int_{\mathcal{L}_e(G)\times K^c} \mathbf{1}_{K_r^c}(\ell_1)\, d\pi_n(\ell_1, \ell_2) \\
&\leqslant \int_{\mathcal{L}_e(G)\times\mathcal{L}_e(G)} \frac{d_L(\ell_1, \ell_2)^2}{r^2}\, d\pi_n(\ell_1, \ell_2) + \sigma_0\big(K^c\big) \\
&\leqslant R^2/r^2 + \varepsilon.
\end{aligned}$$

Taking r large enough, we get $\mu_n(K_r^c) \leqslant 2\varepsilon$, $\forall n \in \mathbb{N}$. This means that the family $(\mu_n)_{n\in\mathbb{N}}$ is tight; therefore $(\pi_n)_{n\in\mathbb{N}}$ is also tight. By Prokhorov theorem, there exists a subsequence (n_m) such that (μ_{n_m}) converges weakly to some $\mu \in \mathcal{P}(\mathcal{L}_e(G))$ and π_{n_m} converges weakly to some $\pi \in \mathcal{P}(\mathcal{L}_e(G) \times \mathcal{L}_e(G))$. It is easy to check that $\pi \in C(\mu, \sigma_0)$. The lower semi-continuity of d_L implies

$$\int\limits_{\mathcal{L}_e(G)\times\mathcal{L}_e(G)} d_L(\ell_1, \ell_2)^2\, d\pi(\ell_1, \ell_2) \leqslant \liminf_{m\to\infty} \int\limits_{\mathcal{L}_e(G)\times\mathcal{L}_e(G)} d_L(\ell_1, \ell_2)^2\, d\pi_{n_m}(\ell_1, \ell_2),$$

which yields

$$W_2(\mu, \sigma_0) \leqslant \liminf_{m\to\infty} W_2(\mu_{n_m}, \sigma_0) \leqslant R.$$

So $\mu \in B_R(\sigma_0)$ and $B_R(\sigma_0)$ is compact with respect to the weak convergence topology. □

Proof of Theorem 1.2. The proof will be split into three steps:

Step 1 *(Finite-dimensional approximation).* Take a sequence of finite partitions $(\mathcal{P}_n)_{n\geqslant 1}$ of $[0, 1]$ finer and finer such that $\bigcup_n \mathcal{P}_n$ is dense in $[0, 1]$. To simplify the notations, we write $G^n = G^{\mathcal{P}_n}$, $\nu_n = \nu_{\mathcal{P}_n}$, $\langle,\rangle_n = \langle,\rangle_{\mathcal{P}_n}$, and $\Lambda_n = \Lambda_{\mathcal{P}_n}$. As ν_n is the heat kernel measure on G^n, there exists a positive smooth function $p_n(x)$ such that $d\nu_n(x) = p_n(x)\, dm(x)$, where m denotes the Haar measure on G^n. We will make abuse of the notation m to denote Haar measures on different G^n. Set $\mu_0^{(n)} = \mu_0 \circ \Lambda_n^{-1} \in \mathcal{P}(G^n)$. Then for every $f \in C_b^\infty(G^n)$,

$$\begin{aligned}\int\limits_{G^n} f(x)\, d\mu_0^{(n)}(x) &= \int\limits_{\mathcal{L}_e(G)} f\big(\Lambda_n(\ell)\big)\, d\mu_0(\ell) = \int\limits_{\mathcal{L}_e(G)} f\big(\Lambda_n(\ell)\big) \frac{d\mu_0}{d\nu}(\ell)\, d\nu(\ell) \\ &= \int\limits_{G^n} f(x)\mathbb{E}_\nu\left[\frac{d\mu_0}{d\nu}(\ell) \middle| \Lambda_n(\ell) = x\right] d\nu_n(x).\end{aligned}$$

It follows that $\mu_0^{(n)}$ is absolutely continuous with respect to ν_n and

$$\frac{d\mu_0^{(n)}}{d\nu_n}(x) = \mathbb{E}_\nu\left[\frac{d\mu_0}{d\nu}(\ell) \middle| \Lambda_n(\ell) = x\right], \quad \nu_n\text{-a.e. } x \in G^n. \tag{3.2}$$

By Jensen's inequality and the convexity of function $s \mapsto s\log s$, we get

$$\mathbf{Ent}_{\nu_n}\big(\mu_0^{(n)}\big) \leqslant \mathbf{Ent}_\nu(\mu_0). \tag{3.3}$$

We will use the following result in finite dimension

Theorem 3.3. *Let G be a compact Lie group, whose Lie algebra $\mathcal{G}$ is endowed with a metric $\langle,\rangle_{\mathcal{G}}$ which is not necessarily* Ad_G*-invariant. For any $V \in C^2(G)$ and $\rho_0 \in \mathcal{P}(G)$ a probability measure on G such that* $\mathbf{Ent}_m(\rho_0) < +\infty$ *where m is the normalized Haar measure on G, then there is a unique solution $\{\rho_t \in \mathcal{P}(G),\ t \in [0, 1]\}$ to the Fokker–Planck equation*

$$\begin{aligned}&\int\limits_{[0,1]\times G} \big[\alpha'(t) f(x) + \alpha(t)\big(\Delta_G f(x) - \big\langle\nabla f(x), \nabla V(x)\big\rangle\big)\big]\, d\rho_t(x)\, dt \\ &\quad = \alpha(0) \int\limits_G f(x)\, d\rho_0(x),\end{aligned} \tag{3.4}$$

for any $\alpha \in C_c^\infty([0,1[)$ *and* $f \in C^2(G)$. *Moreover,*

$$\phi(\rho_t) \leqslant \phi(\rho_0), \quad \textit{and} \quad W_2^2(\rho_t, \rho_s) \leqslant 6(t-s)\big(\phi(\rho_0) - c_0\big) \tag{3.5}$$

where $c_0 = \inf_{\rho \in \mathcal{P}(G)} \phi(\rho)$ *and*

$$\phi(\rho) = \mathbf{Ent}_m(\rho) + \int_G V(x)\, d\rho(x). \tag{3.6}$$

Let ν_G be the heat measure on G associated to the Laplacian defined by the metric $\langle,\rangle_{\mathcal{G}}$, and $d\nu_G = p\, dm$. Take $V = -\log p$; then for any $d\rho = q\, d\nu_G = qp\, dm$,

$$\phi(\rho) = \int_G qp \log(qp)\, dm - \int_G (\log p)\, qp\, dm = \mathbf{Ent}_{\nu_G}(\rho) \geqslant 0.$$

In this case, the relation (3.5) reads as

$$W_2^2(\rho_t, \rho_s) \leqslant 6(t-s)\, \mathbf{Ent}_{\nu_G}(\rho_0). \tag{3.7}$$

Now we use Theorem 3.3 for the case $(G^n, \mathcal{G}^n)$ equipped with the metric defined by (2.4) to get a curve $t \to \mu_t^{(n)}$ on $\mathcal{P}(G^n)$ with density $\rho_n(t,x)$ with respect to the Haar measure m, which satisfies the equation

$$\int_0^1 \int_{G^n} \big[\alpha'(t) f + \alpha(t)\big(\Delta_n f + \langle \nabla f, \nabla \log p_n \rangle_n\big)\big]\, d\mu_t^{(n)}\, dt = \alpha(0) \int_{G^n} f\, d\mu_0^{(n)}, \tag{3.8}$$

for any $\alpha \in C_c^\infty([0,1[)$ and $f \in C^\infty(G^n)$, where Δ_n denotes the Laplacian operator on $(G^n, \langle,\rangle_n)$. It holds that

$$\Delta_n f + \langle \nabla f, \nabla \log p_n \rangle_n = -\operatorname{div}_{\nu_n}(\nabla f), \quad f \in C^\infty\big(G^n\big). \tag{3.9}$$

Therefore, Eq. (3.8) can be rewritten in the form

$$\int_{[0,1]\times G^n} \big[\alpha'(t) f(x) - \alpha(t) \operatorname{div}_{\nu_n}(\nabla f)\big]\, d\mu_t^{(n)}\, dt = \alpha(0) \int_{G^n} f(x)\, d\mu_0^{(n)}. \tag{3.10}$$

Again by Theorem 3.3,

$$\mathbf{Ent}_{\nu_n}\big(\mu_t^{(n)}\big) \leqslant \mathbf{Ent}_{\nu_n}\big(\mu_0^{(n)}\big) \leqslant \mathbf{Ent}_\nu(\mu_0). \tag{3.11}$$

Using the disintegration $\nu(d\ell) = \int_{G^n} \nu(d\ell | \Lambda_n(\ell) = x)\, d\nu_n(x)$, we define $\widetilde{\mu}_t^{(n)} \in \mathcal{P}(\mathcal{L}_e(G))$ by

$$\int_{\mathcal{L}_e(G)} F(\ell) \widetilde{\mu}_t^{(n)}(d\ell) = \int_{G^n} \int_{\mathcal{L}_e(G)} F(\ell) \nu\big(d\ell \big| \Lambda_n(\ell) = x\big)\, d\mu_t^{(n)}(x), \quad \forall F \in \mathbf{Cylin}\big(\mathcal{L}_e(G)\big). \tag{3.12}$$

We have, according to (3.11),

$$\mathbf{Ent}_\nu\big(\widetilde{\mu}_t^{(n)}\big) = \mathbf{Ent}_{\nu_n}\big(\mu_t^{(n)}\big) \leqslant \mathbf{Ent}_\nu(\mu_0). \tag{3.13}$$

By the transportation cost inequality on $\mathcal{L}_e(G)$ [22, Theorem 3.5]:

$$W_2(\rho\nu,\nu)^2 \leqslant C_0\,\mathbf{Ent}_\nu(\rho\nu),$$

where C_0 is a constant related to the lower bound of the Ricci tensor on $\mathcal{L}_e(G)$, and combining with (3.13), we know that $\forall n \geqslant 1$, $\forall t \in [0,1]$, $\widetilde{\mu}_t^{(n)}$ are all in a closed ball of radius $R>0$

$$B_R(\nu) = \bigl\{\mu \in \mathcal{P}\bigl(\mathcal{L}_e(G)\bigr);\ W_2(\mu,\nu) \leqslant R\bigr\}.$$

Thanks to Proposition 3.2, $B_R(\nu)$ is weakly compact. Thus, for each fixed $t \in \mathbb{Q}\cap[0,1]$, there exists a subsequence (n_k) such that $\widetilde{\mu}_t^{(n_k)}$ converges weakly to a probability measure μ_t. Through the diagonal program, we can choose a subsequence of (n_k), denoted again by n_k for simplicity of notation, such that for each $t \in \mathbb{Q}\cap[0,1]$, $\widetilde{\mu}_t^{(n_k)}$ converges weakly to some $\mu_t \in \mathcal{P}(\mathcal{L}_e(G))$.

Step 2 *(Extension of $(\mu)_{t\in\mathbb{Q}}$).* We will extend t to the full interval $[0,1]$; to this end, we will establish the uniform continuity of $t \to \mu_t$ on $\mathbb{Q}$. We claim that, for each pair $t,s \in \mathbb{Q}\cap[0,1]$,

$$W_2(\mu_t,\mu_s)^2 \leqslant 6|t-s|\,\mathbf{Ent}_\nu(\mu_0). \tag{3.14}$$

In fact, for an optimal transport plan $\pi_{n_k} \in C(\mu_t^{(n_k)},\mu_s^{(n_k)})$, we define

$$\widetilde{\pi}_{n_k}\bigl(d\ell,d\ell'\bigr) = \pi_{n_k}(dx,dy)\nu\bigl(d\ell\big|\Lambda_{n_k}(\ell)=x\bigr)\nu\bigl(d\ell'\big|\Lambda_{n_k}\bigl(\ell'\bigr)=y\bigr). \tag{3.15}$$

Then $\widetilde{\pi}_{n_k} \in C(\widetilde{\mu}_t^{(n_k)},\widetilde{\mu}_s^{(n_k)})$. Since $(\widetilde{\mu}_t^{(n_k)})_k$ and $(\widetilde{\mu}_s^{(n_k)})_k$ are tight, $(\widetilde{\pi}_{n_k})_k$ is also tight. Up to a subsequence, $\widetilde{\pi}_{n_k}$ converges weakly to a probability measure π and $\pi \in C(\mu_t,\mu_s)$. We write $d_n(x,y)$ instead of $d_{\mathcal{P}_n}(x,y)$ for any $x,y \in G^n$. Let Λ_n^m be the projection map from G^m to G^n when $m>n$. The definition of $\widetilde{\pi}_{n_m}$ yields that for $n_m > n_k$

$$\int\limits_{\mathcal{L}_e(G)\times\mathcal{L}_e(G)} d_{n_k}\bigl(\Lambda_{n_k}\ell,\Lambda_{n_k}\ell'\bigr)^2\,d\widetilde{\pi}_{n_m}\bigl(\ell,\ell'\bigr) = \int\limits_{G^{n_m}\times G^{n_m}} d_{n_k}\bigl(\Lambda_{n_k}^{n_m}x,\Lambda_{n_k}^{n_m}y\bigr)^2\,d\pi_{n_m}(x,y). \tag{3.16}$$

By continuity of function $(\ell,\ell') \mapsto d_{n_k}(\Lambda_{n_k}\ell,\Lambda_{n_k}\ell')$, we have

$$\int\limits_{\mathcal{L}_e(G)\times\mathcal{L}_e(G)} d_{n_k}\bigl(\Lambda_{n_k}\ell,\Lambda_{n_k}\ell'\bigr)^2\,d\pi\bigl(\ell,\ell'\bigr) \leqslant \liminf_{m\to+\infty} \int\limits_{\mathcal{L}_e(G)\times\mathcal{L}_e(G)} d_{n_k}\bigl(\Lambda_{n_k}\ell,\Lambda_{n_k}\ell'\bigr)^2\,d\widetilde{\pi}_{n_m}\bigl(\ell,\ell'\bigr), \tag{3.17}$$

which is equal to, by (3.16),

$$\liminf_{m\to+\infty} \int\limits_{G^{n_m}\times G^{n_m}} d_{n_k}\bigl(\Lambda_{n_k}^{n_m}x,\Lambda_{n_k}^{n_m}y\bigr)^2\,d\pi_{n_m}(x,y);$$

this last term is dominated by, according to Proposition 2.1,

$$\liminf_{m\to+\infty} \int\limits_{G^{n_m}\times G^{n_m}} d_{n_m}(x,y)^2\,d\pi_{n_m}(x,y) = \liminf_{m\to+\infty} W_2\bigl(\mu_t^{(n_m)},\mu_s^{(n_m)}\bigr)^2 \leqslant 6|t-s|\,\mathbf{Ent}_\nu(\mu_0).$$

Now using again Proposition 2.1 and letting $k\to+\infty$ in (3.17), we get (3.14). $\square$

Now for each $t_0 \in [0,1] \setminus \mathbb{Q}$, we define

$$\mu_{t_0} = \lim_{t \to t_0;\, t \in \mathbb{Q}} \mu_t \quad \text{weakly.}$$

By compactness (see Proposition 3.2) and uniform continuity (see (3.14)) of $t \to \mu_t$, the above extension is well defined. Moreover the inequality (3.14) holds for all $t, s \in [0,1]$. Now by lower semi-continuity of $\mu \to \mathbf{Ent}_\nu(\mu)$ (see [1,24]), the relation (3.13) yields first for $t \in \mathbb{Q}$,

$$\mathbf{Ent}_\nu(\mu_t) \leqslant \mathbf{Ent}_\nu(\mu_0).$$

Again the lower semi-continuity of entropy leads, for $t \in [0,1] \setminus \mathbb{Q}$,

$$\mathbf{Ent}_\nu(\mu_t) \leqslant \liminf_{s \to t,\, s \in \mathbb{Q}} \mathbf{Ent}_\nu(\mu_s) \leqslant \mathbf{Ent}_\nu(\mu_0). \tag{3.18}$$

Recall that (n_k) is the subsequence chosen in the first step such that $\lim_{k\to+\infty} \widetilde{\mu}_t^{(n_k)} = \mu_t$ weakly for all $t \in \mathbb{Q} \cap [0,1]$. Then for each $N \in \mathbb{N}$, we have

$$\lim_{k \to +\infty} (\varLambda_N)_* \widetilde{\mu}_t^{(n_k)} = (\varLambda_N)_* \mu_t \quad \text{weakly, } \forall t \in [0,1]. \tag{3.19}$$

In fact, we remark first that for $t \in \mathbb{Q} \cap [0,1]$ and $n_k \geqslant N$,

$$(\varLambda_N)_* \widetilde{\mu}_t^{(n_k)} = (\varLambda_N)_* \mu_t^{(n_k)}.$$

Secondly, for any smooth function f on G^N, we have

$$\begin{aligned}
&\left| \int_{G^{n_k}} f\big(\varLambda_N(x)\big)\, d\mu_t^{(n_k)} - \int_{G^{n_k}} f\big(\varLambda_N(y)\big)\, d\mu_s^{(n_k)} \right| \\
&\quad = \left| \int_{G^{n_k} \times G^{n_k}} \big[f\big(\varLambda_N(x)\big) - f\big(\varLambda_N(y)\big)\big]\, d\pi_{t,s}^{(n_k)}(x,y) \right| \\
&\quad \leqslant \int_{G^{n_k} \times G^{n_k}} \big| f\big(\varLambda_N(x)\big) - f\big(\varLambda_N(y)\big)\big|\, d\pi_{t,s}^{(n_k)}(x,y)
\end{aligned}$$

where $\pi_{t,s}^{(n_k)}$ is an optimal transport plan in $C(\mu_t^{(n_k)}, \mu_s^{(n_k)})$. By Proposition 2.1,

$$\big| f\big(\varLambda_N(x)\big) - f\big(\varLambda_N(y)\big)\big| \leqslant \|\nabla f\|_\infty\, d_N\big(\varLambda_N(x), \varLambda_N(y)\big) \leqslant \|\nabla f\|_\infty\, d_{n_k}(x,y).$$

So there is a constant $C > 0$ such that

$$\begin{aligned}
&\left| \int_{G^{n_k}} f\big(\varLambda_N(x)\big)\, d\mu_t^{(n_k)} - \int_{G^{n_k}} f\big(\varLambda_N(y)\big)\, d\mu_s^{(n_k)} \right| \\
&\quad \leqslant C W_2\big(\mu_t^{(n_k)}, \mu_s^{(n_k)}\big) \leqslant \sqrt{|t-s|\, \mathbf{Ent}_\nu(\mu_0)}.
\end{aligned}$$

Combining these two points, we obtain (3.19). □

Step 3 *(Solution of Fokker–Planck equation).* In Step 1, we saw that the family $\{\widetilde{\mu}_t^{(n)};\ t \in [0,1],\ n \geqslant 1\}$ is tight. Hence $\{\widetilde{\mu}_t^{(n)}(d\ell) \times dt;\ n \geqslant 1\}$ is a tight family of probability measures on $\mathcal{L}_e(G) \times [0,1]$. Then there exists a subsequence (n_k) such that $\widetilde{\mu}_t^{(n_k)}(d\ell) \times dt$ converges weakly to a probability measure $\bar{\mu}(d\ell, dt)$ on $\mathcal{L}_e(G) \times [0,1]$. Let $\rho_{n_k}(t,\ell)$ denote the density of $\widetilde{\mu}_t^{(n_k)}$ with respect to ν, then, according to (3.13),

$$\int_0^1 \int_{\mathcal{L}_e(G)} \rho_{n_k}(t,\ell) \log \rho_{n_k}(t,\ell)\, d\nu(\ell)\, dt = \int_0^1 \mathbf{Ent}_\nu\big(\mu_t^{(n)}\big)\, dt \leqslant \mathbf{Ent}_\nu(\mu_0) < +\infty.$$

Hence, letting $k \to +\infty$, the lower semi-continuity of relative entropy yields that $\bar{\mu}(d\ell, dt)$ is absolutely continuous with respect to $\nu(d\ell) \times dt$. Furthermore, if $\bar{\rho}(t,\ell)$ denotes its density, then it holds

$$\int_0^1 \int_{\mathcal{L}_e(G)} \bar{\rho}(t,\ell) \log \bar{\rho}(t,\ell)\, \nu(d\ell)\, dt \leqslant \mathbf{Ent}_\nu(\mu_0). \tag{3.20}$$

It follows that for a.e. $t \in [0,1]$, $\mathbf{Ent}_\nu(\bar{\rho}(t,\cdot)) < +\infty$.

Now for any $\alpha \in C_c^\infty([0,1[)$ and any $F \in \mathbf{Cylin}(\mathcal{L}_e(G))$ in the form $F = f \circ \Lambda_N$ for some $N \in \mathbb{N}$ and $f \in C^\infty(G^N)$, we have for $n_k > N$,

$$\int_{\mathcal{L}_e(G)\times[0,1]} \alpha'(t) F(\ell)\, d\widetilde{\mu}_t^{(n_k)}(\ell)\, dt = \int_{G^{n_k}\times[0,1]} \alpha'(t) f\big(\Lambda_N^{n_k}(x)\big)\, d\mu_t^{(n_k)}(x)\, dt. \tag{3.21}$$

By (3.10), the right-hand side of (3.21) is

$$\int_{G^{n_k}\times[0,1]} \alpha(t)\ \mathrm{div}_{n_k}\big(\nabla\big(f \circ \Lambda_N^{(n_k)}\big)\big)(x)\, d\mu_t^{(n_k)}(x)\, dt + \alpha(0) \int_{G^n} f(x)\, d\mu_0^{(n_k)}.$$

Now by (2.12), the first term in above expression comes to,

$$\begin{aligned} &\int_{G^{n_k}\times[0,1]} \alpha(t) \mathbb{E}_\nu^{\mathcal{B}_{n_k}}\big(\Delta^{DM} F\big)(x)\, d\mu_t^{(n_k)}(x)\, dt \\ &\quad = \int_{\mathcal{L}_e(G)\times[0,1]} \alpha(t)\, \Delta^{DM} F\, d\widetilde{\mu}_t^{(n_k)}(\ell)\, dt \\ &\quad \to \int_{\mathcal{L}_e(G)\times[0,1]} \alpha(t)\, \Delta^{DM} F\, d\bar{\mu}(\ell,t) \quad \text{as } k \to +\infty. \end{aligned}$$

Now as $k \to +\infty$, the left-hand side of (3.21) goes to

$$\int_{\mathcal{L}_e(G)\times[0,1]} \alpha'(t) F(\ell)\, d\bar{\mu}(\ell,t).$$

It follows that

$$\begin{aligned} &\int_{\mathcal{L}_e(G)\times[0,1]} \alpha'(t) F(\ell)\, d\bar{\mu}(\ell,t) \\ &\quad = \int_{\mathcal{L}_e(G)\times[0,1]} \alpha(t)\, \Delta^{DM} F\, d\bar{\mu}(\ell,t) + \alpha(0) \int_{\mathcal{L}_e(G)} F\, d\mu_0. \end{aligned} \tag{3.22}$$

On the other hand, for any $\alpha \in C_c^\infty(]0,1[)$ and $F \in \mathbf{Cylin}(\mathcal{L}_e(G))$ in the form $F = f \circ \Lambda_N$ for some $N \in \mathbb{N}$ and $f \in C^\infty(G^N)$, it holds

$$\int_0^1 \int_{\mathcal{L}_e(G)} \alpha(t) F(\ell)\, d\mu_t(\ell)\, dt = \lim_{k\to+\infty} \int_{[0,1]\times\mathcal{L}_e(G)} \alpha(t) f\big(\Lambda_N(\ell)\big)\, d\widetilde{\mu}_t^{(n_k)}\, dt$$
$$= \int_{\mathcal{L}_e(G)\times[0,1]} \alpha(t) F(\ell)\, d\bar{\mu}(t,\ell). \tag{3.23}$$

Therefore,

$$\mu_t(d\ell) \times dt = \bar{\mu}(dt, d\ell), \tag{3.24}$$

and $\mu_t(d\ell) \times dt$ also satisfies the Fokker–Planck equation, which provides an absolutely continuous solution of the Fokker–Planck equation. The proof of this theorem has been completed. □

4. Proof of Theorem 3.3

In this section, the potential V in the definition of the functional ϕ in (3.10) will be $C^2(G)$:

$$\phi(\rho) = \mathbf{Ent}_m(\rho) + \int_G V(x)\, d\rho(x).$$

Let Z be a vector field on G, we denote by T_t the flow associated to Z defined by

$$\frac{dT_t(x)}{dt} = Z\big(T_t(x)\big), \qquad T_0(x) = x.$$

For each $t \in \mathbb{R}$, $x \to T_t(x)$ is a diffeomorphism over G; it is well known [9] that the push forward measure $(T_t)_* m$ has a density h_t with respect to the Haar measure m:

$$h_t(x) = e^{-\int_0^t \operatorname{div}(Z)(T_{-s}(x))\, ds}. \tag{4.1}$$

Proposition 4.1. *Let ρ be a probability measure such that $\mathbf{Ent}_m(\rho) < +\infty$. Then*

$$\frac{d}{dt}\bigg|_{t=0} \phi\big((T_t)_*\rho\big) = -\int_G \big[\operatorname{div}(Z) - \langle \nabla Z, \nabla V\rangle\big]\, d\rho. \tag{4.2}$$

Proof. Let $d\rho = f\, dm$. Then we have $d(T_t)_*\rho = f(T_{-t}) h_t\, dm$, and

$$\mathbf{Ent}_m\big((T_t)_*\rho\big) = \mathbf{Ent}_m(\rho) + \int_G \log\big(h_t(T_t)\big)\, d\rho.$$

Using (4.1) and taking the derivative with respect to t and at $t = 0$, we get

$$\frac{d}{dt}\bigg|_{t=0} \mathbf{Ent}_m\big((T_t)_*\rho\big) = -\int_G \operatorname{div}(Z)(x)\, d\rho(x).$$

On the other side,

$$\frac{d}{dt}\bigg|_{t=0}\int_G V(x)\,d(T_t)_*\rho(x) = \frac{d}{dt}\bigg|_{t=0}\int_G V\big(T_t(x)\big)\,d\rho(x) = \int_G \langle\nabla V, Z\rangle\,d\rho.$$

Combining these last two terms, we get (4.2). □

Following the De Giogri's scheme, explained in [1, pp. 291–292], we have to establish a variational principle. In the case of Riemannian manifolds, a quite general situation has been studied in [26]. In what follows, we shall use McCann's explicit formula for Monge optimal transport maps, which will furnish direct links between two terms in Fokker–Planck equations. We refer to [25] for related topics. Let's first recall McCann's result [19] (see [6] for specific treatments for Lie groups case).

Theorem 4.2. *Let ρ_1, ρ_2 be two probability measure on G. Assume that ρ_1 has a density with respect to the Haar measure m on G; then the Monge optimal transport map $\mathcal{T}$, which pushes ρ_1 forward to ρ_2 has the expression*

$$\mathcal{T}(x) = \exp_x \nabla\Phi(x), \tag{4.3}$$

where $\exp_x(tv)$ denotes the Riemannian geodesic σ, which connects x and $y := \exp_x(v)$ such that $\dot\sigma(0) = v$.

Lemma 4.3. *Let ρ a probability on G such that $\phi(\rho)$ is finite and $\tau > 0$; then there exists a unique ρ_τ of finite entropy $\mathbf{Ent}_m(\rho_\tau) < +\infty$, which is the minimizing point of*

$$\left\{\frac{W_2^2(\rho,\sigma)}{2\tau} + \phi(\sigma);\ \sigma \in \mathcal{P}(G)\right\}. \tag{4.4}$$

Moreover, for any $\psi \in C^2(G)$ and $\nu_t := (T_t)_*\rho_\tau$, where T_t is the flow associated to $\nabla\psi$:

$$\frac{dT_t}{dt} = \nabla\psi(T_t), \qquad T_0(x) = x,$$

it holds

$$\frac{d}{dt}\bigg|_{t=0}\phi(\nu_t) = \int_G \left\langle \frac{\nabla\Phi_\tau}{\tau}, \nabla\psi\right\rangle d\rho_\tau, \tag{4.5}$$

where $\nabla\Phi_\tau$ is given in (4.3).

Proof. The existence of minimizer of

$$\alpha := \inf\left\{\frac{W_2^2(\rho,\sigma)}{2\tau} + \phi(\sigma);\ \sigma \in \mathcal{P}(G)\right\}$$

is obvious due to the fact $\mathcal{P}(G)$ is compact and ϕ is lower semi-continuous with respect to the weak topology on $\mathcal{P}(G)$. If there exist two different minimizers ρ_1 and ρ_2, then for each $\lambda \in (0,1)$, define $\rho_\lambda = (1-\lambda)\rho_1 + \lambda\rho_2 \in \mathcal{P}(G)$. We get

$$\begin{aligned}\phi(\rho_\lambda) &= \mathbf{Ent}\big((1-\lambda)\rho_1 + \lambda\rho_2\big) + (1-\lambda)\int_G V\,d\rho_1 + \lambda\int_G V\,d\rho_2 \\ &< (1-\lambda)\phi(\rho_1) + \lambda\phi(\rho_2),\end{aligned}$$

this last strict inequality is due to the strict convexity of function $s \mapsto s \log s$. Take $\pi_{1,0} \in \mathcal{C}(\rho_1, \rho)$ and $\pi_{2,0} \in \mathcal{C}(\rho_2, \rho)$ two optimal transport plans, then $\pi_{\lambda,0} := (1-\lambda)\pi_{1,0} + \lambda\pi_{2,0} \in \mathcal{C}(\rho_\lambda, \rho)$ and

$$W_2^2(\rho_\lambda, \rho) \leqslant \int_{G\times G} d(x, y)^2 \, d\pi_{\lambda,0}(x, y) = (1-\lambda) W_2^2(\rho_1, \rho) + \lambda W_2^2(\rho_2, \rho).$$

Combining above two estimates, we get $W_2^2(\mu_\lambda, \mu)/(2\tau) + \phi(\mu_\lambda) < \alpha$, which is a contradiction. So we obtain the uniqueness for ρ_τ.

By the minimizing property of ρ_τ, we have

$$\frac{W_2^2(\rho_\tau, \rho)}{2\tau} + \phi(\rho_\tau) \leqslant \phi(\rho) < +\infty,$$

which implies that $\phi(\rho_\tau) \leqslant \phi(\rho)$ and ρ_τ is absolutely continuous w.r.t. m. Using again the minimizing property of ρ_τ, we get

$$\phi(\nu_t) - \phi(\rho_\tau) \geqslant \frac{1}{2\tau}\bigl(W_2^2(\rho, \rho_\tau) - W_2^2(\rho, \nu_t)\bigr). \tag{4.6}$$

The map $x \to \exp_x(\nabla\Phi_\tau(x))$ is the Monge optimal transport map which sends ρ_τ to ρ. We have $(\exp_x \nabla\Phi_\tau \times T_t)_*\rho_\tau \in C(\rho, \nu_t)$, and

$$W_2^2(\rho, \nu_t) \leqslant \int_G d_G^2\bigl(\exp_x\bigl(\nabla\Phi_\tau(x)\bigr), T_t(x)\bigr) \, d\rho_\tau(x). \tag{4.7}$$

Set

$$\eta(t, x) = d_G\bigl(\exp_x\bigl(t\psi(x)\bigr), T_t(x)\bigr).$$

It is obvious that

$$\lim_{t\downarrow 0} \frac{\eta(t, x)}{t} = 0 \quad \text{and} \quad \sup_{(t,x)\in[0,1]\times G} \bigl(\eta(t, x)/t\bigr) < +\infty. \tag{4.8}$$

We have, by triangular inequality,

$$d_G\bigl(\exp_x\bigl(\nabla\Phi_\tau(x)\bigr), T_t(x)\bigr) \leqslant d_G\bigl(\exp_x\bigl(\nabla\Phi_\tau(x)\bigr), \exp_x\bigl(t\nabla\psi(x)\bigr)\bigr) + \eta(t, x).$$

Since G is compact, the diameter D of G is finite; therefore

$$d_G^2\bigl(\exp_x\bigl(\nabla\Phi_\tau(x)\bigr), T_t(x)\bigr) \leqslant d_G^2\bigl(\exp_x\bigl(\nabla\Phi_\tau(x)\bigr), \exp_x\bigl(t\nabla\psi(x)\bigr)\bigr) + 2D\eta(t, x) + \eta(t, x)^2.$$

Combining (4.6), (4.7) and above relation, we have

$$\begin{aligned}&\phi(\nu_t) - \phi(\rho_\tau) \\ &\quad \geqslant \frac{1}{2\tau}\left\{\int_G \bigl[d_G^2\bigl(\exp_x\bigl(\nabla\Phi_\tau(x)\bigr), x\bigr) - d_G^2\bigl(\exp_x\bigl(\nabla\Phi_\tau(x)\bigr), \exp_x\bigl(t\nabla\psi(x)\bigr)\bigr)\bigr] \, d\rho_\tau - \varepsilon(t)\right\}\end{aligned} \tag{4.9}$$

where $\varepsilon(t) = \int_G (2D\eta(t, x) + \eta(t, x)^2) \, d\rho_\tau(x)$. By (4.8) and Lebesgue dominated convergence theorem, $\lim_{t\to 0}(\varepsilon(t)/t) = 0$. To complete the proof of (4.5), we will use the following result due to R. McCann [19], p. 10.

Proposition 4.4. *Let (M, g) be a smooth manifold. Suppose $\sigma : [0, 1] \to M$ has minimal length among piecewise C^1 curves joining $y = \sigma(0)$ to $x = \sigma(1)$, parameterized with constant speed, that is, $\sigma(t) = \exp_y(tu)$ with some $u \in T_yM$. Then $z \to \chi(z) := d_M^2(z, y)/2$ has supergradient $\dot\sigma(1)$ at x, that is*

$$\chi\big(\exp_x(v)\big) \leqslant \chi(x) + \langle v, \dot\sigma(1)\rangle + o\big(|v|\big) \quad \textit{for any } v \in T_xM.$$

We denote $\tilde\sigma(s) = \exp_x(s\nabla\Phi_\tau(x))$; then $\tilde\sigma$ is a minimizing geodesic which send x to $y := \exp_x(\nabla\Phi_\tau(x))$. Now by reversing the time, $\sigma(s) = \tilde\sigma(1-s)$; then σ is a minimizing geodesic sending y to x, and $\dot\sigma(1) = -\nabla\Phi_\tau(x)$. Using Proposition 4.4, we get

$$\begin{aligned}&d_G^2\big(\exp_x\big(\nabla\Phi_\tau(x)\big), x\big)/2 - d_G^2\big(\exp_x\big(\nabla\Phi_\tau(x)\big), \exp_x\big(t\nabla\psi(x)\big)\big)/2\\ &\quad \geqslant \big\langle \nabla\Phi_\tau(x), t\nabla\psi(x)\big\rangle + o\big(|t|\big).\end{aligned}$$

Inserting this inequality in (4.9), we get

$$\lim_{t\downarrow 0}\frac{\phi(\nu_t)-\phi(\rho_\tau)}{t} \geqslant \frac{1}{\tau}\int_G \big\langle \nabla\Phi_\tau(x), \nabla\psi(x)\big\rangle\, d\rho_\tau(x) \geqslant \lim_{t\uparrow 0}\frac{\phi(\nu_t)-\phi(\rho_\tau)}{t}.$$

Proposition 4.1 saying that $\lim_{t\to 0}\frac{\phi(\nu_t)-\phi(\rho_\tau)}{t}$ exists, so we get the equality (4.5). □

Proof of Theorem 3.3. In order to construct the solution to Eq. (3.4). For each time step $\tau > 0$, define inductively $\rho_\tau^{(0)} = \rho_0$ and for $k \geqslant 1$,

$$\rho_\tau^{(k)} = \text{minimizer of } \inf_{\sigma\in\mathcal{P}(G)}\left\{\frac{W_2^2(\rho_\tau^{(k-1)}, \sigma)}{2\tau} + \phi(\sigma)\right\}.$$

Since

$$\frac{W_2^2(\rho_\tau^{(k)}, \rho_\tau^{(k-1)})}{2\tau} + \phi\big(\rho_\tau^{(k)}\big) \leqslant \phi\big(\rho_\tau^{(k-1)}\big),$$

we deduce that

$$\phi\big(\rho_\tau^{(k)}\big) \leqslant \phi\big(\rho_\tau^{(k-1)}\big) \leqslant \cdots \leqslant \phi(\rho_0) < +\infty, \tag{4.10}$$

and

$$\sum_{k=m+1}^{n}\frac{W_2^2(\rho_\tau^{(k)}, \rho_\tau^{(k-1)})}{2\tau} \leqslant \phi\big(\rho_\tau^{(m)}\big) - \phi\big(\rho_\tau^{(n)}\big), \quad n > m. \tag{4.11}$$

Therefore the triangle inequality, combining with Cauchy–Schwarz inequality yields

$$W_2^2\big(\rho_\tau^{(n)}, \rho_\tau^{(m)}\big) \leqslant 2(n-m)\tau\left[\phi\big(\rho_\tau^{(m)}\big) - \phi\big(\rho_\tau^{(n)}\big)\right]. \tag{4.12}$$

Let

$$c_0 = \inf_{\sigma\in\mathcal{P}(G)}\phi(\sigma). \tag{4.13}$$

It is obvious that $c_0 \in \mathbb{R}$ is bounded below. Define $\{\bar\rho_t; t \in [0, 1]\}$ by

$$\bar\rho_\tau(k\tau) = \rho_\tau^{(k)}, \quad k = 1, \ldots, N,$$

where $N = [1/\tau]$ the integral part of $1/\tau$ and for $t, s \in](k-1)\tau, k\tau]$, we connect $\rho_\tau^{(k-1)}$ and $\rho_\tau^{(k)}$ by constant speed geodesic: for $t \in [(k-1)\tau, k\tau]$, $\bar{\rho_\tau}(t)$ is given by

$$\bar{\rho_\tau}(t) = \left[\exp_x\left(\frac{k\tau - t}{\tau}\nabla\Phi_k(x)\right)\right]_* \rho_\tau^{(k)}, \tag{4.14}$$

where $x \to \exp_x \nabla\Phi_k(x)$ is the Monge optimal map pushing $\rho_\tau^{(k)}$ forward $\rho_\tau^{(k-1)}$. We have $W_2(\bar{\rho}_\tau(t), \bar{\rho}_\tau(s)) = |t-s| W_2(\rho_\tau^{(k-1)}, \rho_\tau^{(k)})$ for $t, s \in](k-1)\tau, k\tau]$. For $t = n\tau$ and $s = m\tau$, the relation (4.12) reads as

$$W_2\big(\bar{\rho}_\tau(t), \bar{\rho}_\tau(s)\big) \leqslant \sqrt{|t-s|}\sqrt{2\big(\phi(\rho_0) - c_0\big)}. \tag{4.15}$$

For $t \in](n-1)\tau, n\tau]$ and $s \in](m-1)\tau, m\tau]$ with $m < n$ we have

$$\begin{aligned}
&W_2\big(\bar{\rho}_\tau(t), \bar{\rho}_\tau(s)\big) \\
&\quad\leqslant W_2\big(\bar{\rho}_\tau(t), \bar{\rho}_\tau\big((n-1)\tau\big)\big) + W_2\big(\bar{\rho}_\tau\big((n-1)\tau\big), \bar{\rho}_\tau(m\tau)\big) + W_2\big(\bar{\rho}_\tau(m\tau), \bar{\rho}_\tau(s)\big) \\
&\quad\leqslant \big(t - (n-1)\tau\big) W_2\big(\rho_\tau^{(n-1)}, \rho_\tau^{(n)}\big) + \sqrt{2\big(\phi(\rho_0) - c_0\big)}\sqrt{\big|(n-1)\tau - m\tau\big|} \\
&\qquad + (m\tau - s) W_2\big(\rho_\tau^{(m-1)}, \rho_\tau^{(m)}\big) \\
&\quad\leqslant \sqrt{6\big(\phi(\rho_0) - c_0\big)}\sqrt{|t-s|}.
\end{aligned}$$

So $\bar{\rho}_\tau(\cdot)$ takes its value in $C([0,1], \mathcal{P}(G))$, and by Arzela–Ascoli theorem, there exists $(\rho(t))_{t\in[0,1]}$ such that, up to a subsequence, $\bar{\rho}_\tau(t)$ converges to $\rho(t)$ uniformly with respect to $t \in [0,1]$ as $\tau \to 0^+$. Furthermore, we have

$$W_2\big(\rho(t), \rho(s)\big) \leqslant \sqrt{6\big(\phi(\rho_0) - c_0\big)}\sqrt{|t-s|}. \tag{4.16}$$

Now consider $\tau = 2^{-q}$. By (4.10), for any $p \geqslant q$ and $k = 0, \ldots, 2^q$, we have $\phi(\bar{\rho}_{2^{-p}}(k2^{-q})) \leqslant \phi(\rho_0)$. As $p \to +\infty$, $\bar{\rho}_{2^{-p}}(k2^{-q})$ converges weakly to $\rho(k2^{-q})$; the lower semi-continuity [1] of ϕ yields

$$\phi\big(\rho\big(k2^{-q}\big)\big) \leqslant \phi(\rho_0).$$

Since the set $D = \{k2^{-q};\ q \geqslant 1,\ k = 0, \ldots, 2^q\}$ is dense in $[0,1]$; for each $t \in [0,1]$, we can find $s_n \in D$ such that $s_n \to t$. So the above inequality leads

$$\phi\big(\rho(t)\big) \leqslant \phi(\rho_0), \quad t \in [0,1]. \tag{4.17}$$

In the sequel, we will show that $(\rho_t)_{t\in[0,1]}$ is a solution of (3.4). Let $\alpha \in C_c^\infty([0,1[)$ and $f \in C^\infty(G)$; set

$$I_k = \int_{(k-1)\tau}^{k\tau} \alpha'(t) \int_G f(x)\, d\bar{\rho}_\tau(t)\, dt - \int_{(k-1)\tau}^{k\tau} \alpha'(t) \int_G f(x)\, d\rho_\tau^{(k)}\, dt. \tag{4.18}$$

Since G is compact, it holds that

$$f\left(\exp_x\left[\frac{k\tau - t}{\tau}\nabla\Phi_k(x)\right]\right) - f(x) = O\left(\frac{k\tau - t}{\tau}\big|\nabla\Phi_k(x)\big|\right), \tag{4.19}$$

where O is uniform with respect to k. Notice that

$$\int_G \left|\nabla\Phi_k(x)\right| d\bar\rho_\tau^{(k)}(x) \leqslant \left(\int_G \left|\nabla\Phi_k(x)\right|^2 d\bar\rho_\tau^{(k)}(x)\right)^{1/2} = W_2\left(\rho_\tau^{(k-1)}, \rho_\tau^k\right).$$

So according to (4.18) and (4.19), we get, for some constant $C_{\alpha,f} > 0$,

$$\sum_{k=1}^N |I_k| \leqslant C_{\alpha,f}\,\tau\, W_2\left(\rho_0, \bar\rho_\tau^{(N)}\right) \leqslant C_{\alpha,f}\,\tau^{3/2}\sqrt{2\left(\phi(\rho_0)-c_0\right)}. \tag{4.20}$$

Now set

$$J_k = \int_{(k-1)\tau}^{k\tau} \alpha'(t)\int_G f(x)\,d\rho_\tau^{(k-1)}\,dt. \tag{4.21}$$

We have

$$\begin{aligned}\sum_{k=1}^N J_k &= \sum_{k=1}^N\left[\left(\alpha(k\tau)-\alpha\left((k-1)\tau\right)\right)\int_G f(x)\,d\rho_\tau^{(k-1)}(x)\right]\\ &= \alpha(N\tau)\int_G f(x)\,d\rho_\tau^{(N-1)} - \alpha(0)\int_G f\,d\rho_0\\ &\quad + \sum_{k=1}^{N-1}\alpha(k\tau)\left[\int_G f(x)\,d\rho_\tau^{(k-1)}(x) - \int_G f(x)\,d\rho_\tau^{(k)}(x)\right]\\ &= \alpha(N\tau)\int_G f(x)\,d\rho_\tau^{(N-1)} - \alpha(0)\int_G f\,d\rho_0\\ &\quad + \sum_{k=1}^{N-1}\alpha(k\tau)\int_G \left[f\left(\exp_x\left(\nabla\Phi_k(x)\right)\right) - f(x)\right]d\rho_\tau^{(k)}(x).\end{aligned}$$

Notice that

$$f\left(\exp_x\left(\nabla\Phi_k(x)\right)\right) - f(x) = \left\langle\nabla f(x), \nabla\Phi_k(x)\right\rangle + O\left(\left|\nabla\Phi_k(x)\right|^2\right),$$

and $\sum_{k=1}^N \int_G |\nabla\Phi_k(x)|^2\,d\rho_\tau^{(k)} = \sum_{k=1}^N W_2^2(\rho_\tau^{(k-1)}, \rho_\tau^{(k)})$ which is less, by (4.11), than $2\tau\cdot(\phi(\rho_0)-c_0)$. So

$$\begin{aligned}\sum_{k=1}^N J_k &= \alpha(N\tau)\int_G f(x)\,d\rho_\tau^{(N-1)} - \alpha(0)\int_G f\,d\rho_0\\ &\quad + \sum_{k=1}^{N-1}\alpha(k\tau)\tau\int_G \left\langle\nabla f(x), \nabla\Phi_k(x)/\tau\right\rangle d\rho_\tau^{(k)} + \varepsilon(\tau),\end{aligned}$$

with $\lim_{\tau\to 0}\varepsilon(\tau) = 0$. Now by Proposition 4.1 and Lemma 4.3,

$$\int_G \left\langle\nabla f(x), \nabla\Phi_k(x)/\tau\right\rangle d\rho_\tau^{(k)}(x) = -\int_G \left[\Delta_G f - \langle f, V\rangle\right] d\rho_\tau^{(k)}.$$

Therefore the term

$$\sum_{k=1}^{N-1} \alpha(k\tau)\tau \int_G \langle \nabla f(x), \nabla \Phi_k(x)/\tau \rangle \, d\rho_\tau^{(k)}$$

$$= -\sum_{k=1}^{N-1} \int_{(k-1)\tau}^{k\tau} \alpha(k\tau)\, ds \int_G \big[\Delta_G f - \langle f, V\rangle\big]\, d\rho_\tau^{(k)}$$

which goes to, as $\tau \to 0$

$$-\int_{[0,1]\times G} \alpha(t)\big[\Delta_G f(x) - \langle f(x), V(x)\rangle\big]\, d\rho(t)\, dt.$$

On the other hand

$$\lim_{\tau\to 0} \int_{[0,1]\times G} \alpha'(t) f(x)\, d\bar{\rho}_\tau(t)\, dt = \int_{[0,1]\times G} \alpha'(t) f(x)\, d\rho(t)\, dt.$$

Therefore the above calculations lead (3.4). □

Acknowledgement

The second named author would like to thank Professor Theo Sturm for discussions and encouragements during his stay at the Bonn University, where an essential part of this work has been done.

References

[1] L. Ambrosio, N. Gigli, G. Savaré, Gradient Flows in Metric Spaces and in the Space of Probability Measures, Lectures in Mathematics, ETH/Birkhäuser Verlag, Zürich/Basel, 2005.
[2] L. Ambrosio, G. Savaré, L. Zambotti, Existence and stability for Fokker–Planck equations with log-concave reference measure, Probab. Theory Related Fields 145 (2009) 517–564.
[3] B. Driver, Integration by parts and quasi-invariance for heat measures on loop groups, J. Funct. Anal. 149 (1997) 470–547.
[4] B. Driver, T. Lohrenz, Logarithmic Sobolev inequalities for pinned loop groups, J. Funct. Anal. 140 (1996) 381–448.
[5] S. Fang, J. Shao, Transportation cost inequalities on path and loop groups, J. Funct. Anal. 218 (2005) 293–317.
[6] S. Fang, J. Shao, Optimal transport maps for Monge–Kantorovich problem on loop groups, J. Funct. Anal. 248 (2007) 225–257.
[7] S. Fang, J. Shao, K.T. Sturm, Wasserstein space over the Wiener space, Probab. Theory Related Fields 146 (2010) 535–565.
[8] D. Feyel, A.S. Üstünel, Monge–Kantorovitch measure transportation and Monge–Ampère equation on Wiener space, Probab. Theory Related Fields 128 (2004) 347–385.
[9] H. Kunita, Stochastic Flows and Stochastic Differential Equations, Cambridge University Press, 1990.
[10] R. Jordan, D. Kinderlehrer, F. Otto, The variational formulation of the Fokker–Planck equation, SIAM J. Math. Anal. 29 (1998) 1–17.
[11] J. Lott, C. Villani, Ricci curvature for metric-measure spaces via optimal transport, Ann. of Math. 169 (2009) 903–991.
[12] P. Malliavin, Stochastic calculus of variation and hypoelliptic operator, in: K. Itô (Ed.), Proceedings of the International Symposium on Stochastic Differential Equations, Kyoto, 1976, Tokyo, 1976.
[13] M.P. Malliavin, P. Malliavin, Integration on loop groups, I. Quasi-invariance measures, J. Funct. Anal. 93 (1990) 207–237.

[14] H. Airault, P. Malliavin, Integration on loop groups, II. Heat equation for the Wiener measure, J. Funct. Anal. 104 (1992) 71–109.
[15] P. Malliavin, Integration on loop groups III. Asymptotic Peter–Weyl orthogonality, J. Funct. Anal. 108 (1992) 13–46.
[16] P. Malliavin, Hypoellipticity in infinite dimension, in: M. Pinsky (Ed.), Diffusion Processes and Related Problems in Analysis, Birkhäuser, Boston, MA, pp. 17–23.
[17] Z.M. Ma, M. Röckner, Introduction to the Theory of (Nonsymmetric) Dirichlet Forms, Universitext, Springer-Verlag, Berlin, 1992.
[18] R.J. McCann, A convexity principle for interacting gases, Adv. Math. 128 (1997) 153–179.
[19] R.J. McCann, Polar factorization of maps on Riemannian manifolds, Geom. Funct. Anal. 11 (2001) 589–608.
[20] F. Otto, The geometry of dissipative evolution equations: the porous medium equation, Comm. Partial Differential Equations 26 (2001) 101–174.
[21] F. Otto, C. Villani, Generalization of an inequality by Talagrand, and links with the logarithmic Sobolev inequality, J. Funct. Anal. 173 (2000) 361–400.
[22] J. Shao, Hamilton–Jacobi semi-groups in infinite dimensional spaces, Bull. Sci. Math. 130 (2006) 720–738.
[23] K.T. Sturm, Convex functionals of probability measures and nonlinear diffusions on manifolds, J. Math. Pures Appl. 84 (2005) 149–168.
[24] K.T. Sturm, On the geometry of metric measure spaces, Acta Math. 196 (2006) 65–131.
[25] C. Villani, Optimal Transport, Old and New, Grundlehren der mathematischen Wissenschaften, vol. 338, Springer, Berlin, Heidelberg, 2009.
[26] X. Zhang, Variational approximation for Fokker–Planck equation on Riemannian manifold, Probab. Theory Related Fields 137 (2007) 519–539.